

Meeting of the Group of Experts on Business Registers

Online 26-29 September 2022

Session 2: The use of administrative data, big data, and other data sources

**Integrating trade register data in the Statistical Business Register – recent developments in Germany**

Lorna Syme, Robin Lorenz

Federal Statistical Office, Germany

**Abstract**

The Statistics Register Act (StatRegG) allows the statistical offices in Germany use of information from the national trade register to maintain the Statistical Business Register.

Some regional offices have been processing this information, with varying intensity, for years. In 2019 a working group was formed to create a harmonized concept across all regional offices.

Through regular processing of the information in the trade registry data, core information about a legal unit such as address and legal form can be updated. Since the trade registration number is a unique identifier, ensuring it is up-to-date also allows for linkages between the Statistical Business Register and other data sources.

Notifications that are published by the trade registry can also be used to maintain the Statistical Business Register. These notifications are published daily and announce for example firm foundings and closures, or name changes in existing legal units. Processing this information will increase the currency and quality of the information in the Statistical Business Register.

The challenge of processing these notifications is extracting the relevant information from unstructured text. Techniques have been developed in this area to identify various structural events such as foundings or changes in name or address. In the future other information could be extracted from these notifications, for example the description of business activities which could then be coded to determine a firm's NACE classification.

This paper outlines the advantages and challenges of a systematic, harmonized processing of information from the trade registry to maintain the Statistical Business Register.

## 1. Introduction

The German Statistical Business Register (hereinafter simply Business Register) is a key tool used in structural analyses of the German economy. The quality and accuracy of official statistics published in Germany is dependent on the quality and accuracy of the information contained in the Business Register.

The Business Register is maintained using a number of administrative sources, the main sources being monthly data from the Federal Employment Agency and from the financial administration. Additional sources include data from the Chamber of Crafts, the Deutsche Bundesbank (central bank) and the EuroGroups Register. Information on enterprise groups is supplied by a commercial data provider and information gained from surveys also flows back into the Business Register.

The statistical offices of the Federation and the Länder (federal states) are charged with maintaining the Business Register and continually strive to increase the quality and accuracy of the information it contains. This includes searching for new sources of reliable information that can be integrated into the Business Register.

One such source is the national trade register which since 2007 has been maintained in electronic form and through the Statistics Register Act is made available for use in the maintenance of the Business Register. The trade register consists of two parts: index data for all firms registered in the trade register and ongoing notifications of changes to firms in the trade register.

All statistical offices of the Länder have access to the data from the trade register. Four of these offices – Bavaria, Baden-Württemberg, Hesse and North Rhine-Westphalia – already systematically process this information to maintain those legal units in the Business Register under their respective mandates. In 2019 a working group was formed to develop a concept to be adopted across all statistical offices.

This concept must take into account the large volume of information available from the trade register as well as the heterogeneity of available resources and personnel across the statistical offices of the Länder. The concept therefore stipulates a mandatory annual processing of the index data of the trade register but also allows for an optional processing of the ongoing notifications.

The following sections describe the data of the trade register and discuss the benefits of its integration in the Business Register as well as the associated challenges.

## 2. The trade register

The trade register, register of cooperatives and register of partnerships (hereinafter trade register) contains information on all firms, cooperatives and partnerships registered in Germany. While not covering all legal forms (for example certain partnerships) and not including the self-employed, the trade register does contain all limited companies, public limited companies and sole traders and therefore provides coverage of a large number of economically active firms in Germany.

Index data on these registered firms is received by the statistical offices of the Federation and the Länder annually while daily notifications are bundled and received monthly.

## 2.1 Index data

Variables in the index data include firm name, address, legal form, date of registration, status of registration (current or closed<sup>1</sup>), register section, registration court, and the register number assigned by the registration court. These final three variables create the unique identifier for a particular entry in the trade register. Register section, registration court and the register number are variables maintained in the Business Register. This will facilitate the initial linkage between the trade register and Business Register that will be required in the first year of the processing. This identifier is publicly available and widely used, also allowing for linkages to other data sources.

*Table 1: Trade register units and Business Register units in the trade register*

<b>Register type</b>	<b>Number of registrations in the trade register (February 2022)</b>	<b>Number of legal units in the Business Register (31st December 2021, preliminary)</b>
Trade Register, Section A	466,339	430,527
Trade Register, Section B	1,464,443	1,379,545
Register of cooperatives	9,110	8,060
Register of partnerships	17,152	12,488
<b>Total</b>	<b>1,957,044</b>	<b>1,830,580</b>

Table 1 shows the number of entries in the trade register and the number of legal units in the Business Register which are assigned to the various sections. The distribution between the sections is similar: the majority of entries are in section B of the trade register which covers public limited companies, limited partnerships and limited companies. Section A covers sole traders and general partnerships and makes up the next largest portion of entries. In total the index data include information on 1,957,044 firms, cooperatives and partnerships while in the Business Register 1,830,530 legal units are registered in the trade, cooperative or partnership register.

<sup>1</sup> The current registrations are relevant for the initial processing of the trade register. Thereafter firms which are no longer active in the subsequent year can be identified by their 'closed' status in the index data. Information of this closure also appears in the monthly notification data. The processing of these notifications, however, is optional. Those statistical offices of the Länder who do not process these notifications will still be able to identify the closure with the next delivery of the index data.

Caution should be taken when interpreting the difference in number of units between the two registers. While the higher number of entries in the trade register may indicate that the trade register contains legal units which are currently unknown to the Business Register, it may simply be that these units are contained in the Business Register but the information on register section, registration court and register number is missing. One goal of the processing of the trade register data is to ensure the best possible coverage of this important identifier.

## 2.2 Notifications

The trade register also publishes daily notifications (German: *Bekanntmachungen*). These are short texts announcing new registrations, closures and mergers, as well as changes in firm name and changes of address. These notifications are bundled and made available electronically on a monthly basis. The notification data includes a variable with date of notification, the notification text in full, plus the register section, registration court and the register number allowing a linkage to the index data.

The relevant information for the processing of these notifications must first be extracted from the notification text. This is done by searching the text for key words or phrases which indicate certain relevant business events. Figure 1 shows an example of a notification text announcing a change in firm address. In this case the string ‚Change of address‘ is extracted and coded. In total over 30 relevant events are extracted and coded from the notification texts. Table 2 shows an excerpt of the more relevant of these events.

*Figure 1: Example of notification text*

HRB 34483: CRANIO Investment, Leipzig, Kuhnaustraße 1, 04279 Leipzig. Change of address: Thomaskirchhof 20, 04109 Leipzig.  
16.02.2022, Registration court Leipzig HRB 34483 (translated from German)

*Table 2: Relevant events extracted from notification texts (excerpt)*

<b>Event</b> <i>(translated from German)</i>	<b>Notifications</b> <i>(February to December 2021)</i>
Registration ended	85,173
Change of business address	46,457
Change in registered office	58,381
Change to firm	32,829
Change to corporate purpose	29,127
Merger	16,969
Notification of deletion	8,498

## 3. Use of trade register data for maintaining the Business Register

The trade register contains information of firms in Germany which are legally required to be registered and to keep the register informed of changes to firm name,

firm address, legal form, etc. Because of this the trade register offers a reliable source for maintaining the core information of legal units in the Business Register. Once an initial link to the trade register has been established via the variables register section, registration court and the register number, the subsequent updating of this core information can take place more or less automatically. This is especially beneficial given the large number of datasets in the trade register.

Using the trade register data also increases the coverage of the variables register section, registration court and the register number for registered firms in the Business Register. These variables form the basis for the legal entity identifier number (LEID) for the EuroGroups Register and is a widely-used, publicly available identifier, allowing for linkages to other data sources.

By processing trade register notifications, information about changes of address or firm name can be incorporated in a timely fashion: a change can be identified and updated in the Business Register within a month of notification. This increases the accuracy of the information in the Business Register and reduces the number of non-responses in surveys due to incorrect addresses.

Integrating the trade register data into the Business Register is not without its challenges. While the register section, registration court and the register number allow for the unique identification of an entry in the trade register, this does not necessarily equate to a unique identification of a firm. When a firm relocates to a different jurisdiction it is required to register in the court of that jurisdiction and receives a new register number. From the point of view of the trade register this is a new dataset, for the Business Register this is the same legal unit. Furthermore, when a registration is closed, the responsible registration court is free to re-issue the same register number to a new firm. In this case the incorrect information would be linked to the legal unit in the Business Register.

The processing of the notifications also poses challenges as relevant events must be extracted and coded from unstructured texts. There are 146 registration courts in Germany and no single standardised reporting. All possible variants must be manually added to the programming and continually updated. The notification text may also be very lengthy and may contain announcements of multiple events bundled into one text field. The identification and extraction of the individual events in this case is technically complex. The coded events must then also be prioritised in the subsequent processing.

#### 4. Conclusion

Integrating trade register data into the Business Register allows for timely identification of changes to firms and structural changes, and increases the accuracy of the Business Register. Processing the index data annually ensures good coverage of the trade register identifier (register section, registration court, register number), facilitating linkages to other data sources.

The techniques currently used to extract relevant events from the notifications could be expanded to exploit other available information. For example when a firm registers with the trade register for the first time it is required to declare its firm

activities. A change to firm activity must also be declared. These text descriptions may help with the assignment of firms into NACE classifications given their activities, with the caveat that the descriptions must first be extracted from unstructured text.

The notifications could also be exploited to allow for the timely identification of new firms. It is theoretically possible to identify and integrate a new firm into the Business Register within a month of the firm's founding. While not part of the current concept, this could be integrated into the processing in the future.

## References

Schneider, Volker. *Digitalisierung in der amtlichen Statistik – Nutzung von Verwaltungsdaten*. In: Statistisches Monatsheft Baden-Württemberg. Issue 5/2019, pp. 28 – 31.

[https://www.statistik-bw.de/Service/Veroeff/Monatshefte/PDF/Beitrag19\\_05\\_05.pdf](https://www.statistik-bw.de/Service/Veroeff/Monatshefte/PDF/Beitrag19_05_05.pdf)  
(accessed 13.09.2022)

Stephan, Frank/ Rödel, Raimund. *Nutzung von Bekanntmachungen des elektronischen Handelsregisters zur Qualitätssicherung im Statistischen Unternehmensregister Bayern*. In: Bayern in Zahlen. Issue 10/2019, pp. 639 – 640. [https://www.statistik.bayern.de/mam/produkte/biz/z1000g\\_201910.pdf](https://www.statistik.bayern.de/mam/produkte/biz/z1000g_201910.pdf)  
(accessed 13.09.2022)

Von Eschwege, Katja. *Neuerungen im Statistischen Unternehmensregister: Auswertungskonzept, Relevanzschwellen und weitere Quellen*. In WISTA. Issue 5/2021, pp. 86 – 98.

[https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2021/05/neuerungen-unternehmensregister-052021.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2021/05/neuerungen-unternehmensregister-052021.pdf?__blob=publicationFile)  
(accessed 13.09.2022)